Clinical Linguistics & Phonetics, Sept–Nov 2005; 19(6/7): 503–514

Taylor & Francis
Taylor & Francis Group

Techniques for field application of lingual ultrasound imaging

BRYAN GICK[1,2], SONYA BIRD[3], & IAN WILSON[1]

[1]*Department of Linguistics, University of British Columbia,* [2]*Haskins Laboratories,* [3]*Department of Linguistics, University of Victoria*

(Received 4 March 2004; accepted 15 December 2004)

Abstract
Techniques are discussed for using ultrasound for lingual imaging in field-related applications. The greatest challenges we have faced distinguishing the field setting from the laboratory setting are the lack of controlled head/transducer movement, and the related issue of tissue compression. Two experiments are reported. First, a pilot study identifies important factors in controlling head/transducer movement in field settings. Second, an Optotrak/ultrasound study reports the range of head movement in an optimal field-like setting within and across varying phonetic contexts, as well as the effect of tongue tissue compression on tongue image data. Results suggest that with a simple arrangement involving a head rest or surface, a fixed transducer, and careful design and presentation of stimuli, reliable lingual ultrasound data can be collected in the field.

Keywords: *Ultrasound, fieldwork, linguistic phonetics, intervention, validation*

Technological advances have enabled speech researchers and clinicians to measure directly the movements of the vocal tract during speech with ever-increasing detail and accuracy using a wide variety of tools. Ultrasonography has proven particularly useful for accessing new populations of subjects and patients, being safer and less invasive than many other imaging tools, and often readily available in hospitals and clinics. However, there have been very few attempts to extend ultrasound work beyond the laboratory or clinic, excluding a huge pool of potential participants, whether due to age, disability, remoteness, means, or simply inconvenience. Many sources are available that describe techniques that have been used for applying ultrasound technology to speech (see other papers in this issue; also, e.g., Stone, 1997), including one previous paper describing applications for phonetic fieldwork (Gick, 2002). While it is clear that ultrasound offers many advantages over other tools for speech imaging (non-invasiveness, safety, ease of data collection, instant feedback, relatively high sampling rate, portability, etc.), a central issue has persisted in determining methods used for ultrasound research in speech, in or out of the laboratory: the lack of absolute spatial reference in the signal. The present paper describes and evaluates some of the techniques recently used for ultrasound applications in field settings, with particular

Correspondence: Bryan Gick, E270-1866 Main Mall, Vancouver, BC, CANADA V6T 1Z1. Tel: 604 822 4347. Fax: 604 822 9687. E-mail: gick@interchange.ubc.ca

ISSN 0269-9206 print/ISSN 1464-5076 online
DOI: 10.1080/02699200500113590

focus on the success of different field data collection set-ups in controlling the unwanted effects of head movement and lingual floor tissue compression.

When an ultrasound transducer is held against the skin of the neck beneath the chin and above the larynx, the ultra-high frequency sound emitted by its piezoelectric crystals travels upward through the tongue and is reflected back to the transducer from the tongue-air interface. The resulting echo patterns are used to reconstruct a 2-dimensional image of the tongue surface (in any plane), which can then be viewed or recorded from the machine as a video image for analysis or visual feedback. While ultrasound is well suited for measuring the location of soft tissues, it is not able to image through bone or air. The resulting signal thus contains images representing only the distance from the transducer to the tongue surface, but without any stable spatial reference structures such as the palate, other skull bone points, or cervical vertebrae. The only way to reconstruct tongue position relative to the rest of the vocal tract is therefore to control or correct for transducer movement relative to head position. As accurate correction would require cumbersome tracking equipment and advanced reconstruction techniques not available to the typical practitioner, the present paper will focus on methods for limiting (rather than tracking) head movement under normal field conditions.

A variety of different head/transducer restraint devices have been used in previous studies to constrain unwanted movement in laboratory or clinical settings (e.g., Akgul, Kambhamettu, & Stone, 1998; Peng, Jost-Brinkmann, Miethke, & Lin, 2000). However, the extra equipment required for such controls are for the most part too cumbersome, and often too invasive, for field applications. One promising technique involves mounting the ultrasound probe on a helmet worn by the subject, so that the head and transducer move together. This method, currently being developed by the Queen Margaret University College (QMUC) ultrasound group (see Gibbon, 2005), may prove valuable in cases where the helmet arrangement is workable for the subject, but will be more difficult for new researchers to adopt, as it requires specially designed equipment. Under normal field conditions, and in the absence of specialized equipment, there will invariably be some unwanted movement. In such cases, simple physical controls can be imposed on the head or transducer. While various portable alternatives for limiting movement of the head and transducer have been experimented with in previous field studies, the advantages of one method over another have been largely anecdotal. A related issue concerns the effect of submental tissue compression by direct contact with the transducer head. Because of the lack of spatial reference, excessive tissue compression can cause changes in transducer-to-tongue distance that are mistaken for tongue movements. While the effects of tissue compression have been shown to be mitigated in the laboratory using an acoustic standoff (Stone, 2005; Peng et al., 2000), the goal of the present paper is to provide baseline data on the accuracy that can be achieved using only the basic equipment available in a typical field setting.

The present paper first tests which controls normally available in field settings are most successful at limiting head/transducer movement, and second, reports the range of head movement in the resulting configuration, and its effect on lingual ultrasound data. The remainder of this paper considers a number of methods that can be implemented using only a portable ultrasound machine and commonly available equipment, such as a chair, laser pointer, and microphone stand. These methods are first evaluated via a pilot study to determine the most effective arrangement for limiting head/transducer movement in field settings. Second, a follow-up study using ultrasound imaging and Optotrak (optical point-tracking) reports (1) the range of motion of head movement of subjects in an optimal field-like setting, and (2) the magnitude of the effect of tongue tissue compression on tongue image data recorded in this setting.

Experiment 1: Methods

A single-subject experiment was conducted to test for head and transducer movement under varying field conditions. Ultrasound data were collected under three different experimental conditions, combined in five trials. Conditions were (a) hand-held vs. fixed transducer (trials 1–4 vs. 5), (b) unrestricted head movement vs. head resting against a surface (trials 1, 3, 5 vs. 2, 4), and (c) presence vs. absence of visual feedback of transducer movement (trials 1, 2 vs. 3, 4). Movement was tracked using video and measured from markers on the subject's head and the transducer. Of interest in this experiment is the variability over time within a trial in the spatial values of a set of measurement points on both head and transducer, with no movement (minimum variability) as the desideratum.

Subject

The subject who participated in the experiment was one of the researchers: a native speaker of English from Montreal, Canada, 30-years-old. She had not previously been a subject in an ultrasound experiment.

Materials

A Sonosite 180 Plus portable ultrasound machine was used with a Sonosite C15/4-2 MHz MCX transducer (see http://www.sonosite.com). The experiment was recorded at standard video rate (29.97 fps) using two Sony (mini-DV) Handycam Vision DCR-TRV900 (NTSC) digital video recorders. An external Shure SM58 microphone recorded the reading task performed during the experiment as well as the subject's qualitative evaluation of each trial. The output of the microphone and two video recorders were combined using a Videonics MXProDV digital audio/video mixer and the combined signal was recorded onto a JVC SR-VS20 Professional DV recorder. The resulting video was transferred to an Apple Macintosh computer running OS 9.2.2 using Adobe Premiere, v.6.0.1. All measurements were taken from the video with ImageJ v. 1.31 (cross-platform graphics software available as freeware from NIH http://rsb.info.nih.gov/ij/).

Procedures

All data were collected and analysed at the Interdisciplinary Speech Research Laboratory at the University of British Columbia. Figure 1 illustrates the experimental setup. In order to record head and transducer movement, the two video cameras recorded the subject, one from the side view and the other from the front. The subject sat in a solid, straight-backed, wooden chair, approximately 2 meters away from both video cameras. The stimuli to be read were placed on a music stand just to the left of the subject (as close to centre as possible without interfering with the frontal video camera). An external microphone was used to record the sentences being read and the subject's impressions of each trial. A small (pea-sized) amount of ultrasound gel was applied to the head of the transducer prior to the start of each trial to improve skin-transducer contact.

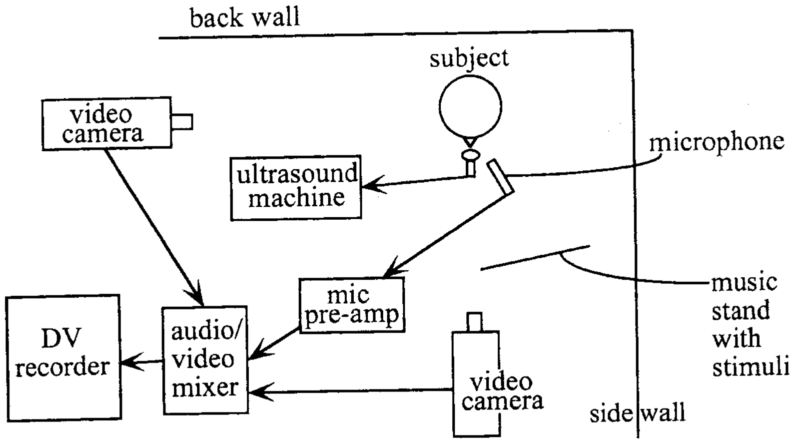

Figure 1. Experiment 1 arrangement.

To facilitate taking measurements, red markers were placed on the transducer (front and side), on the subject's neck, and on glasses worn by the subject (one marker on the right stem and another on the nose bridge). A point on the helix of the subject's right ear was used for a head angle measurement. Figure 2 illustrates the side and front views of the subject and the placement of the markers on the transducer and on her head.

Conditions were (a) hand-held vs. fixed transducer (Trials 1–4 vs. 5), (b) unrestricted head movement vs. head resting against a surface (Trials 1, 3, 5 vs. 2, 4), and (c) presence vs. absence of visual feedback of transducer movement (Trials 1, 2 vs. 3, 4). Experimental trials were thus as follows:

Trial 1. The subject sat on a chair with no restrictions on head movement, and held the transducer herself.

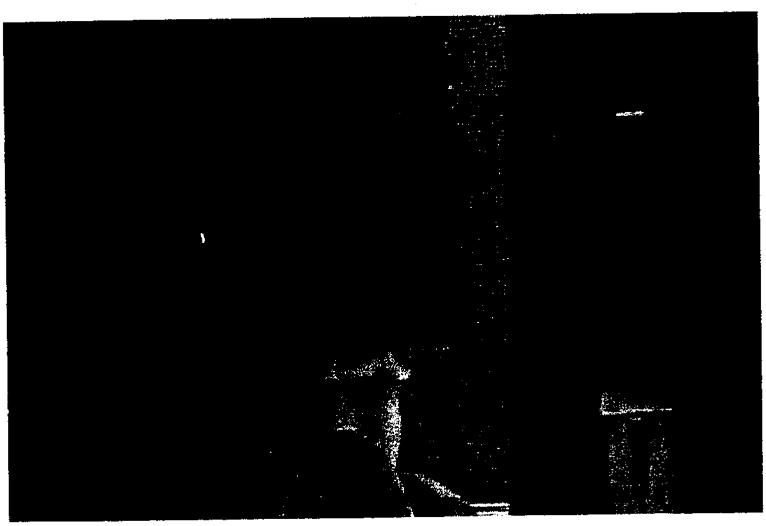

Figure 2. Front and side views of the subject and the markers used to take measurements in Experiment 1.

Trial 2. Identical to Trial 1, except that head movement was restricted by having the subject lean her head against the wall behind her.

Trial 3. As Trial 1, except a laser pointer was attached to the transducer and aimed at a crosshair target in front of the subject, so that she could visually monitor transducer movement herself and attempt to compensate for it throughout the experiment.

Trial 4. As Trial 3, except that the subject's head movement was restricted by leaning her head against the wall behind her (this is the method presented in Gick, 2002).

Trial 5. As Trial 1, except the transducer was attached to a rigid microphone stand on the floor in front of the subject.

All trials for the experiment were collected in one day, with a minimum of 5 minutes of rest between trials. Data were recorded while the subject was continuously reading English sentences for 3 minutes from a stimuli list consisting of short sentences containing a wide range of English sounds. The day before the experiment, the subject was given two practice trials to test the equipment and setup. On the day of the experiment, the subject was given another practice trial to assimilate her to the set-up.

Analysis

All measurements were made in pixels and converted to millimeters based on a calibration ruler that was placed in the videos adjacent to each plane of interest: one on the side-view video at the points of the glasses, one on the side-view video at the transducer, and one on the front-view video at the points of the glasses and transducer. The conversion was between 2.14 and 2.16 pixels per millimeter depending on the plane. The measurement error introduced was therefore approximately .5 mm. Still frames were extracted from the video recordings at 10-second intervals throughout the 3-minute test period using Adobe Premiere, giving 19 still frames for each trial. Thus, measures totaled 19 samples of each of the 16 measurements for each of the five trials. In total, 1520 measurements were taken. All measurements are described in Table I. In addition, a number of calculations were made based on the measurements in Table I. These are given in Table II.

In addition to quantitative results of head and transducer movement, the subject was asked to provide her impressions of each trial in terms of comfort level and general difficulty of the required task. Results are therefore of two types: (a) qualitative observations from the subject and (b) quantitative observations based on the measurements taken.

Experiment 1: Results

As described above, the experimental trials were chosen to illustrate the differences in head and transducer movement under the following conditions: (a) hand-held vs. fixed transducer (Trials 1–4 vs. 5), (b) unrestricted head movement vs. head resting against a surface (Trials 1, 3, 5 vs. 2, 4), and (c) presence vs. absence of visual feedback of transducer movement (Trials 1, 2 vs. 3, 4).

Qualitative results

Comfort level was found to be somewhat higher when the subject held the transducer herself (Trials 1–4) than when the transducer was fixed (Trial 5). The fixed transducer

Table I. Measurements used in Experiment 1 analysis

	Measurement	Description
	Side view	
0	Head angle	Angle between lines connecting glasses marker to ear helix to neck marker
1	Head mid vert	Vert dist from right of glasses marker to screen bottom
2	Head front vert	Vert dist from left of glasses marker to screen bottom
3	Head-probe vert	Vert dist from right of glasses marker to transducer centre
4	Head side horiz	Horiz dist from right of glasses marker to left screen edge
5	Probe rear vert	Vert dist from bottom left of transducer marker to screen bottom
6	Probe mid vert	Vert dist from bottom middle of transducer marker to screen bottom
7	Probe front vert	Vert dist from bottom right of transducer marker to screen bottom
8	Probe side horiz	Horiz dist from right of transducer marker to left screen edge
	Front view	
9	Head vert	Vert dist from front glasses marker to screen bottom
10	Head front horiz	Horiz dist from front glasses marker to screen right edge
11	Nose front horiz	Horiz dist from nose tip to screen right edge
12	Probe top horiz	Horiz dist from top of transducer front marker to screen right edge
13	Probe mid horiz	Horiz dist from centre of transducer marker to screen right edge
14	Probe low horiz	Horiz dist from bottom of transducer marker to screen right edge
15	Probe drift horiz	Horiz dist from chin centre to left edge of transducer head

restricted jaw movement somewhat, making speech perceptibly more difficult (this problem can be mitigated by using a transducer with a smaller head), or using an acoustic standoff (Stone, 2005; Peng et al., 2000). Comfort level was also higher when the head was free (Trials 1, 3 and 5) than when movement was restricted (Trials 2 and 4). This was likely due to the fact that restricting head movement consisted of having the subject lean her head against the wall, which made it awkward to perform the reading task (the head rest position can easily be brought to a more comfortable position by attaching a block to the wall behind the subject's head). During the laser pointer condition (Trials 3 and 4) the subject reported that it was difficult to monitor the laser pointer crosshairs while attempting to focus on the stimuli. Finally, across all trials the subject found that it was more difficult to hold the transducer stable under the chin when gel was freshly applied (gel tends to be more slippery when fresh, and is tackier once it begins to dry).

Quantitative results

Standard deviations were calculated for each of the 16 measures and seven calculations given in Tables I and II, to compare variation in head position, transducer position, and head vs. transducer position across the five trials described above. Analyses of variance were then conducted to compare standard deviations of all measures across the five trials.

Table II. Calculations used for Experiment 1

Measurements	Calculation	Description
2-1	Head nod	Head front vert minus Head mid vert
2-6	Head-probe vert	Head front vert minus Probe mid vert
7-5	Probe front-back tilt	Probe front vert minus Probe rear vert
8-4	Head-probe side horiz displacement	Probe side horiz minus Head side horiz
11-10	Head tilt	Nose front horiz minus Head front horiz
13-10	Head-probe front horiz displacement	Probe mid horiz minus Head front horiz
14-12	Probe side-to-side tilt	Probe low horiz minus Probe top horiz

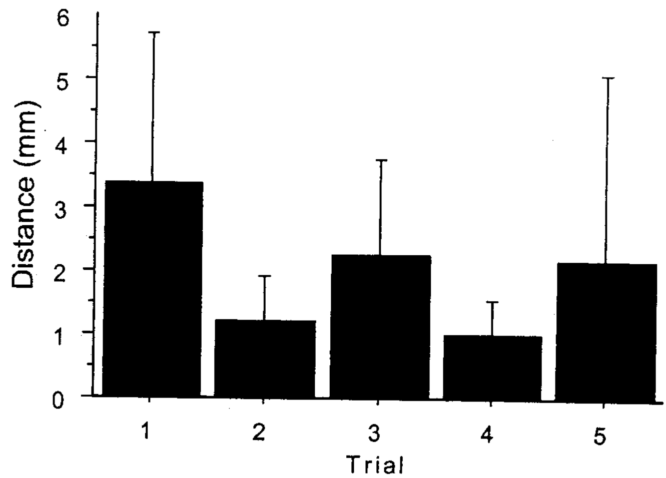

Figure 3. Mean standard deviations of all 23 measures and calculations for Experiment 1, compared by Trial. Error bars show standard deviations from these means.

Overall ANOVA results indicate significant variance in the data ($F\ [4, 110]=6.117$, $p<.001$). Mean standard deviation values are given in Figure 3. Post-hoc analysis (Fisher's PLSD) indicates significant differences ($p<.05$) between the following trials: 1, 2; 1, 3; 1, 4; 1, 5; 3, 4; 4, 5; and near-significant differences for 2, 3 ($p=.0597$) and 2, 5 ($p=.0746$).

When considering only standard deviations in the head vs. probe measures, significant differences were present ($F\ [4, 20]=3.122$, $p<.05$), with Trial 5 showing significantly greater deviations than all other trials (Fisher's PLSD $p<.05$ for all comparisons). This was largely due to very large horizontal drift of the transducer relative to the head during Trial 5 (in which the transducer was fixed, but the head was not).

Experiment 1: Discussion

Results of the pilot validation study suggest that the most important factor under field data collection conditions is to ensure that subjects are provided with a headrest, even if it is only a wall. This not only serves the obvious purpose of limiting head movement, but also limits transducer and head vs. transducer movement, to a mean standard deviation of about 1 mm in the present experiment. In the absence of a headrest, providing visual feedback gives a slight advantage, but it adds no additional advantage when a headrest is used. Another option for visual feedback is to place a mirror directly in front of subjects, so that they can monitor transducer or head movement and compensate for it (Lundstrom & Lundstrom, 1992). However, subjects may feel self-conscious about their speech when in front of a mirror, and have difficulty concentrating on the task at hand. In any event, the present results suggest that visual self-monitoring is likely to be unnecessary where it is possible to use a headrest. Use of a headrest will also greatly reduce the drift relative to a fixed transducer (as in Trial 5).

Experiment 2: Methods

A study was conducted to follow up on the pilot experiment described above, using much more accurate optical tracking techniques to report head movement in an idealized field-like setting involving a completely naive subject and a headrest. Optotrak data were used to measure head position and rotation. In addition, ultrasound data were collected simultaneously to test for significant lingual tissue compression within three phonetic contexts. In the event of lingual tissue compression, we expect to see a positive correlation between head-to-probe distance and tongue-to-probe distance along the probe angle, for a given articulatory target, assuming a relatively stable tongue position for that target (i.e., if both the head and tongue move down towards the probe, a similar displacement should be observed unless tongue tissue is being compressed).

Subject

The subject who participated in this study was a 27-year-old male speaker of Wisconsin English. The subject was paid for his participation and he was unaware of the purpose of the study.

Procedures

The subject was seated in a modified ophthalmic examination chair with his head resting against a fixed headrest (see Figure 4). The subject's head was otherwise unrestrained. Four types of data were simultaneously collected during the experiment: audio data of the speech signal, video data of a frontal view of the subject from the chest up, ultrasound data of tongue movements and Optotrak data of the positions of markers on the head, lips, and ultrasound probe. The audio data were used only for time-aligning the ultrasound and Optotrak data, and the video data were used only for post-monitoring of the experiment. B-mode (2-dimensional) ultrasound data was collected using an Aloka ProSound SSD-5000 ultrasound machine with a UST-9118 endo-vaginal 180° probe. The probe was held in a fixed position against the subject's neck by a rigid mechanical arm, approximating the effect of a microphone stand-mounted transducer in an actual field setting. An Optotrak 3020 infrared tracking system was used to track the movements of eight infrared-emitting diodes (markers). Four markers were attached to a pair of glasses assumed to be fixed relative to the subject's skull. This allowed for rigid body tracking of the head and calculation of the movement of any point assumed to be fixed relative to the head. An additional two markers were attached to the lips, but data from these were not used in this experiment. The final two markers were attached to the ultrasound probe, allowing for the calculation of the angle of the probe. The Optotrak data were collected at 90 Hz and recorded the (x, y, z) coordinates of each marker with greater than .1 mm accuracy. Figure 4 shows the positions of the subject, transducer, headrest, and Optotrak markers.

Stimuli consisted of 30 sentences of the type 'John said "hoo roo" each autumn', where the underlined words always varied but were consistently a person's name, a nonsense two-syllable phrase, and a time, designed to contain a wide range of English vowels and consonants. Immediately prior to data collection, a practice block of 15 sentences was used to help assimilate the subject to the stimuli. The sentences were presented to the subject one at a time on a computer screen using an automated PowerPoint slide show. Each sentence was displayed for 3 seconds followed by a blank screen for 1 second. The computer screen was at eye level, approximately 2 meters in front of the subject, a position known to help limit head movement (Stone, 2005). In addition, prior to ultrasound data collection, about 30 seconds of Optotrak data were collected while the subject rotated his head about the x-axis and y-axis in turn (i.e., the subject was told to turn to the right,

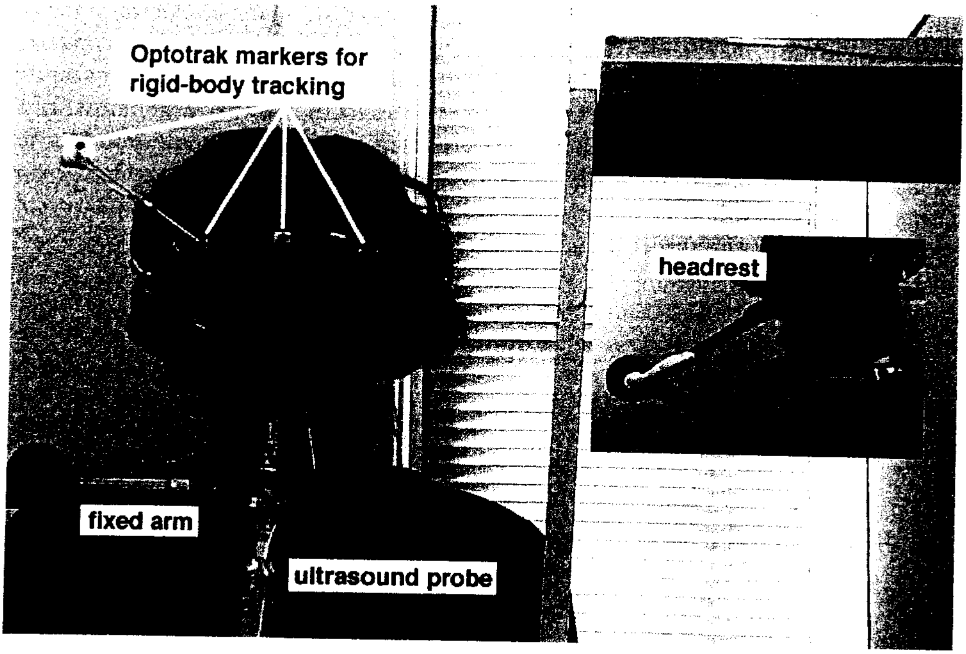

Figure 4. Front and rear views of the subject in Experiment 2, showing the positions of the transducer and headrest, as well as the markers used to take measurements.

centre, left, centre, up, centre, down, centre, and to go through this cycle twice). This enabled the Optotrak software to calculate a pivot point for the subject's head. All rotational data in this experiment are calculated based on that pivot point.

Analysis

Head movement can occur in six dimensions: translation along the x (right-left), y (up-down), and z (front-back) axes, and rotation about each of these 3 axes (yaw (i.e., "nod"), pitch (i.e., "shake"), and roll (i.e., "tilt"), respectively). To determine head movement, the location of a fixed point was calculated approximately where the tongue meets the palate (7 cm inferior and 2.5 cm posterior to the center point on the bridge of the glasses). The x, y, z positions and three rotations of this point were calculated from the positions of the four fixed Optotrak markers attached to the glasses. The resulting coordinates were used to determine overall variation in head movement continuously throughout the trial, as well as variation within and between three phonetic target contexts, [ɛ] in "said", [i] in "each", and inter-utterance rest position.

To measure tongue tissue compression in each of the 3 contexts, [ɛ], [i], and rest, ultrasound-measured positions of the tongue height during these three relatively stable events were compared to Optotrak-measured head positions along the same line, approximately at the palate. Ultrasound data were recorded to DV tape and then captured to a Macintosh G4 computer using Adobe Premiere 6.0. For each of the 30 sentences, three frames were extracted and analyzed: (1) when the tongue body (TB) reached its peak (relative to the angle of the transducer) during [ɛ], (2) when the TB reached its peak during [i], and (3) when the tongue had returned to its speech rest position following the sentence-final word. These frames were opened in Adobe Photoshop 7.0 and the "probe distance" of the tongue (i.e., position of the highest point on the tongue along the line determined by the angle of the transducer) was measured in pixels using the measure tool, then converted to millimeters based on the visual scale in the ultrasound images. The Optotrak head position data at the estimated palate point (see above) were subjected to a rotation into the coordinate system of the ultrasound transducer, extracted at times corresponding to the ultrasound frames, and the "probe distance" was measured. The resulting Optotrak head position measures could thus be compared directly to the ultrasound measures of the tongue within each of the target phonetic contexts.

Table III. Ranges (maximum - minimum) in head position and rotation for Experiment 2—standard deviations in parentheses

Dimension	overall	[i]	[E]	rest
x (right-left)	4.17 mm (.65mm)	2.22 mm (.57mm)	2.58 mm (.66mm)	1.86 mm (.48mm)
y (up-down)	9.40 mm (1.79mm)	2.40 mm (.56mm)	3.47 mm (.88mm)	3.86 mm (.77mm)
z (front-back)	7.14 mm (1.03mm)	3.79 mm (.82mm)	4.14 mm (1.14mm)	2.11 mm (0.57mm)
x rotation/yaw (nod)	6.50° (.87°)	3.18° (.64°)	3.68° (.74°)	3.04° (.54°)
y rotation/pitch (shake)	1.87° (.31°)	.80° (.21°)	.86° (.18°)	.96° (.22°)
z rotation/roll (tilt)	2.99° (.27°)	.95° (.23°)	.79° (.22°)	1.10° (.27°)

Experiment 2: Results

Table III shows the overall variation (range between absolute maximum and minimum values) measured continuously across all phonetic contexts throughout the trial, the variation within the three phonetic contexts ([ɛ], [i], and rest), and the standard deviations, in head position and rotation at the calculated palate location.

Differences in head position and rotation between the three phonetic target contexts ([ɛ], [i], and rest), in each of the six dimensions, were tested using ANOVAs. Only two dimensions (y position (up-down) and x rotation/yaw (nod)) showed both significant variance and mean spatial/rotational differences greater than 1mm/1°: (F [2, 85]=147.27, $p<.0001$ and F [2, 85]=20.761, $p<.0001$, respectively). Mean differences were: for y position ([i]>[ɛ]>rest), −0.93 mm ([ɛ] vs. [i]), 2.35 mm ([ɛ] vs. rest), 3.28 mm ([i] vs. rest); for x rotation ([ɛ]>[i]>rest), 0.43° ([ɛ] vs. [i]), 1.10° ([ɛ] vs. rest), 0.67° ([i] vs. rest). Post-hoc analyses (Fisher's PLSD) indicate that all differences are significant ($p<.05$) between all contexts, in both of these dimensions.

Tongue compression results showed no significant correlations between head position and tongue position along the transducer angle within any of the three phonetic contexts included in this study. Results of two-tailed Pearson's correlations for [ɛ], [i], rest, respectively, are: r (df=28, crit=.361)=−.058; r (df=28, crit=.361)=−.144; r (df=26, crit=.374)=−.263.

Discussion

Results of Experiment 2 suggest that with limited controls on head movement, along with a fixed transducer, a relatively high degree of accuracy can be obtained in the context of a simple experimental set-up. While the overall range of head movement across the wide variety of phonetic targets throughout the trial neared 10 mm along one (y) dimension, the range was greatly reduced when comparing like contexts, with positional variation ranging from 1.86 mm to 4.14 mm, and rotational variation ranging from .79° to 3.68°. Not surprisingly, the greatest overall positional variation was seen along the y (up-down) axis, and the greatest overall rotational variation around the x (nod) axis. These were also the only axes that showed significant variance across phonetic targets. These two measurements, displacement along the y-axis and rotation about the x-axis, actually represent a single effect: rotation about the x-axis. Rotating the fixed palate point, described in the analysis section, 6.5° about the pivot point from its starting point accounts for about 8.6 mm of displacement along the y-axis (close to the total of 9.40 mm overall for this axis). These axes correspond with the vertical displacements of the head known to accompany variations in F0 (Yehia, Kuratate, & Vatikiotis-Bateson, 2002), an observation which itself may partly explain the reduction in variation of head position within like phonetic targets, as these targets also occupied identical sentence positions, and were thus subject to identical effects of intonation and stress. Thus, although it will not eliminate all of the effects of head movement, stimulus design may provide an important indirect tool for partial control of head movement for field ultrasound research.

Results for the differences in head position and rotation across phonetic contexts show that mean values in both the y (up-down) position and x rotation (nod) dimensions were higher for vowel targets [i] and [ɛ] than for rest position. This may be due to a more open jaw position or a stiffer tongue floor for the vowel targets.

Finally, the results for tongue floor tissue compression give no evidence of a correlation between head-to-probe distance as measured with Optotrak and tongue-to-probe distance as measured with ultrasound within any of the three phonetic contexts included in this study (/ɛ/, /i/, and rest). As described above, in the event of lingual tissue compression, we expect to see a positive correlation between head-to-probe distance and tongue-to-probe distance along the probe angle, for a given articulatory target, assuming a relatively stable tongue position for that target (i.e., if both the head and tongue move down towards the probe, a similar displacement should be observed unless tongue tissue is being compressed). Thus, as there was some variation in head position, this variation must either have been small enough not to have had a significant impact, or alternatively, it is possible that as the transducer pushes on the mouth floor, the jaw or the tongue surface is pushed up with it resulting in the observed lack of correlation between the head and tongue positions within a single phonetic context. More work is needed in this area to determine the extent of the perturbation effect of the transducer. If significant tongue tissue compression is found in future studies, then it may be necessary to include an acoustic standoff in the standard ultrasound field kit.

Conclusions

Various field applications for lingual ultrasonography were discussed in this paper, and a number of techniques were evaluated as a step toward identifying best methods for use by practitioners in a typical field setting. The findings of the pilot study suggest that the simple solution of resting the subject's head against a surface during data collection has the most dramatic effect on controlling movement of the head and transducer. The findings of the second study show that a simple arrangement involving a head rest and a fixed transducer, along with stimuli that are well designed and visually carefully placed, can help to reduce head movement. Because the present study found that the greatest head displacement occurred along the y-axis, future innovation resulting in better control of vertical head movement would be particularly effective. Results also suggest that tongue floor stiffness and jaw height may have a measurable effect on head position and rotation. It is not clear from the present results the degree to which tongue tissue compression affects ultrasound measurements.

Regardless of which methods are ultimately found to give the best results in field applications, it is clear that the needs of specific situations require flexibility (e.g., working with children will require different methodological choices from working with the elderly). While ongoing work is needed to validate new and different methods, it is likely that a variety of techniques will continue to compete in future published work using portable ultrasound.

Acknowledgements

The authors wish to recognize the contributions of the many users who have tested and contributed to the systems described in this paper, who can not all be mentioned here. Particular thanks for help with this paper are due to: Fiona Campbell, Jason Chang, Oliver Guenther, Jeremy Perkins, Bosko Radanov, Shaffiq Rahemtulla, and Eric Vatikiotis-Bateson. This work was funded by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada, and a New Opportunities grant from the Canadian Foundation for Innovation, both to the first author, and by National Institutes of Health Grant DC-02717 to Haskins Laboratories.

References

Akgul, Y. S., Kambhamettu, C., & Stone, M. (1998). Automatic motion analysis of the tongue surface from ultrasound image sequences. *Proceedings of The IEEE Workshop on Biomedical Image Analysis* (pp. 126–132). Santa Barbara, CA.

Gibbon (2005).

Gick, B. (2002). The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association, 32,* 113–122.

Lundstrom, F., & Lundstrom, A. (1992). Natural head position as a basis for cephalometric analysis. *American Journal of Orthodontics and Dentofacial Orthopedics, 101,* 244–247.

Peng, C. L., Jost-Brinkmann, P. G., Miethke, R. R., & Lin, C. T. (2000). Ultrasonographic measurement of tongue movement during swallowing. *Journal of Ultrasound in Medicine, 19,* 15–20.

Stone, M. (1997). Laboratory techniques for investigating speech articulation. In W. J. Hardcastle & J. Laver (Eds.), *Handbook of Phonetic Sciences* (Oxford: Blackwell).

Stone, M. (2005). A guide to analyzing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics.*

Yehia, H. C., Kuratate, T., & Vatikiotis-Bateson, E. (2002). Linking facial animation, head motion and speech acoustics. *Journal of Phonetics, 30,* 555–568.